Generalized Variable Range Hopping Near Two-Dimensional Metal-Insulator Transitions

Wenjun Zheng and Yue Yu
Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100080, P.R.China
(February 7, 2008)

In an attempt to understand quantitatively the remarkable discoveries of metal-insulator transitions in two-dimensional systems, we generalize Mott's variable range hopping theory to the situation with strong Coulomb interaction. In our formulation, the Gaussian form is adopted into the expression of the hopping probability, and the effect of Coulomb gap is also considered. After taking account of the newly proposed scaling consideration, we produce the dynamical and localization length exponents, which are consistent with the experiments. We then clarify the physical content of our formulation and explain the universality of the localization length exponent suggested by a series of experiments. We also discuss the general scaling function of both temperature and electrical field on the insulating side of the transition.

The conventional scaling theory for disordered systems holds that without interactions any extent of disorder is sufficient to localize the electrons and no true metallic behavior is possible at $T = 0$ in two dimensions [1]. The inclusion of interactions to this picture proved to be much more difficult, and no satisfactory theoretical picture has emerged in spite of decades of continuous effort. However the belief that all the states are localized at $d = 2$ has remained unchallenged.

Recently, a series of surprising experiments have been conducted on 2D electron gas and hole gas in zero magnetic field, which demonstrated convincing evidences in support of a true metal-insulator transition(MIT) in 2D [2]. In these experiments, The resistivity scales with temperature or electrical field with a single scaling parameter that approaches zero at a critical carrier density n_c. Further more the scaling function of resistivity near the transition possesses a very simple exponential form, which leads to the reflection symmetry across the transition, as emphasized in ref [3]. That is,

$$\rho(\delta n, T) = 1/\rho(-\delta n, T), \tag{1}$$

where $\delta n = (n - n_c)/n_c$, ρ is the resistivity measured in unites of its critical value ρ_c. This symmetry is very similar to some recent experimental results in quantum Hall systems [4], although the mechanism may be quite different [5]. These remarkable discoveries have evoked a wave of theoretical efforts to understand them [6] [7]. Although it is generally agreed that strong interactions play an important role, no specific microscopic mechanism has been able to explain the experiments both qualitatively and quantitatively.

In a recent Letter [7], Dobrosavljevic et al. gave a general scaling consideration of the 2D MITs. In their formulation, they treated the temperature-dependence of the conductance in the quantum critical region. After assuming the MIT occurs at $g = g_c$ where $\beta(g_c) = 0$ and introducing the scaling variable $t = \log(g/g_c)$, they considered the linear approximation to $\beta(t)$ near $t = 0$ and obtained

$$\beta(t) = \frac{dt}{d(\log L)} \approx \frac{t}{\nu} + O(t^2).$$

By integrating from l to L, they got

$$t(L) = t_0(\frac{L}{l})^{\frac{1}{\nu}},$$

where $t_0 = \log(g_0/g_c)$ is given by the value g_0 of the conductance at the microscopic scale l. Then after using $t_0 \approx (g_0 - g_c)/g_c \propto \delta n$, where $\delta n = (n - n_c)/n_c$ and n is the density, they reached

$$g(L) = g_c \exp[A\delta n(L/l)^{\frac{1}{\nu}}].$$

Then representing the length scale L by the temperature as $T \propto L^{-z}$, one finally arrived at

$$g(\delta n, T) = g_c \exp(A\delta n/T^{\frac{1}{\nu z}}). \tag{2}$$

This expression gives the explicit exponential form of the scaling function of the conductance with respect to temperature except for the exponents ν and z to be determined, and the reflection symmetry in eq(1) follows from it directly, which are precisely what has been seen in the experiments. As emphasized by the authors [7], the above relation is only valid inside the critical region, or for $T > T_0 \propto |\delta n|^{\nu z}$, which is also verified experimentally [3]. In what follows we will use this result as a restrictive condition to our general derivation, which is the generalization of Mott's well-known variable range hopping (VRH) theory. Then we will arrive at the explicit scaling function of the resistivity, from which the exponents $\nu = 1.5$ and $z = 1$ can be read off easily, which are consistent with the experiments [8]. We then discuss the physical mechanism of our formulation and explain the universality of the exponent $\nu = 1.5$ [9], which is characteristic of a general two dimensional systems with strong, long-range interactions(like Coulomb interaction). At last, we discuss the general scaling function of both temperature and electrical field.

Before starting with our formulation, let us give an introduction to some background that motivates the present work. In an influential work [10], Mott suggested his celebrated theory of VRH which described an important conducting mechanism in a localized electron system. His central idea is to optimize the hopping probability with respect to the hopping distance R, which gave the temperature-dependence of the conductivity

$$\sigma(T) \propto \exp[-(\frac{T_0}{T})^{\frac{1}{d+1}}],$$

where the dimension $d = 2$ here. This result has received substantial experimental supports in the past decades [11]. Later, Efros et al. made an attempt to include the effect of Coulomb interaction into the VRH theory [12]. They only considered the Coulomb gap correction to the density of states and obtained

$$\sigma(T) \propto \exp[-(\frac{T_0}{T})^{\frac{1}{2}}],$$

This has also been verified in the experiments performed on dilute electron systems where Coulomb interaction dominates the transport properties [13]. Although this relation can produce the right exponent $z = 1$, it is nevertheless inconsistent with the concrete expressions derived from the accurately designed measurements performed recently [2] and fails to explain the general localization length exponent $\nu \approx 1.5$ revealed by more than one experiment.

In the following, we will extend the VRH idea to the situation with strong Coulomb interaction. Our generalization goes beyond the one attempted by Efros et al., and aims at a better understanding of the recently surging discoveries of MITs in two dimensional systems, especially the exponent ν, which seems to be a universal exponent.

The hopping probability for a particle to cover a distance of R is

$$p(R) \propto \exp(-\alpha\frac{R^2}{\xi^2} - \frac{\Delta}{k_B T}), \tag{3}$$

where ξ is the localization length on the insulator side of the MIT, which diverges as $\xi \propto |\frac{n-n_c}{n_c}|^{-\nu}$ when approaching the MIT (ν is the localization length exponent); α is a non-singular parameter, Δ is the energy difference between the initial state and the final state for the hopping particle. The only difference from Mott's equation is the Gaussian form dependence of $p(R)$ on R, which is vital to the final conclusion and contains a profound physics content(see below). If there is no interaction, we simply have

$$\Delta \propto [R^d N(E_F)]^{-1},$$

as it is supposed in Mott's derivation, where $N(E_F)$ is the density of states at the Fermi level E_F. However, for the case of strong Coulomb interaction, the Coulomb gap will depress the density of states near the Fermi surface. Following Efros et al. [14], we can write

$$N(E) \propto |E - E_F|,$$

which means $N(E_F) = 0$. Therefore $N(E_F)$ in the above equation should be substituted by the following integration which counts the number of states available for a R-range-hopping:

$$\begin{aligned} N(E_F) &\to \frac{1}{\Delta E}\int_0^{\Delta E} dE N(E + E_F) \\ &\propto \Delta E \\ &\propto \frac{1}{R}, \end{aligned} \tag{4}$$

where a linear dispersion $\Delta E \propto \Delta k$ has been used, and Δk is the momentum quanta proportional to $1/R$.

Therefore $\Delta \propto 1/R^{d-1}$. This means the effect of Coulomb interaction is to reduce the dimension by 1. Therefore in what follows, we will use the effective dimension $d' = d - 1$.

Then we maximize eq(3) with respect to R, and obtain

$$R_{max} \propto (\frac{\xi^2}{T})^{\frac{1}{d'+2}}.$$

Therefore we finally arrive at

$$\sigma(n, T) \propto p(R_{max}) \propto \exp\{-\frac{(|\Delta n|)^{\frac{2d'\nu}{d'+2}}}{T^{\frac{2}{d'+2}}}\}. \tag{5}$$

This is the central equation of the present paper, which should be contrasted with the relations of Mott and Efros et al..

Taking into account the classical scaling relation

$$\sigma(n, T) = F(\frac{\Delta n}{T^b}),$$

where $F(x)$ is the scaling function, $b = 1/z\nu$, and z is the dynamical exponent, we can easily see

$$z = d' = 1$$

for a two dimensional system. This result is well-known for a Coulomb interaction dominated system [15]. This relation is also verified in many recent experiments on 2D MIT [2].

However, there is more to obtain from eq(5). Since the above single parameter scaling argument has given the following restrictive form [7]

$$\rho(n, T) \propto \exp(-A\frac{\delta n}{T^b}),$$

we have an additional condition to satisfy:

$$\frac{2d'\nu}{d'+2} = 1,$$

which gives

$$\nu = \frac{d' + 2}{2d'} = 1.5.$$

This is very close to the localization length exponent reported widely by a series of experiments [2]. Therefore,

$$\rho(n, T) \propto \exp(\frac{|\Delta n|}{T^{2/3}}). \tag{6}$$

We would like to comment that even if we have adopted the traditional $p \propto \exp(-R/\xi)$ instead of the Gaussian form, we can still arrive at $z = 1$. However the desirable ν can not be obtained in that way. That is to say, the traditional VRH theory is not consistent with the scaling argument presented in ref [7], under the condition that $\nu = 1.5, \ \ z = 1$.

Now let us give some comments on the physics contained in the Gaussian form adopted in eq(3). This form is quite natural and reasonable in physics, because of its direct connection with the eigenstates of the harmonic oscillator(HO), one of the few universal and analytically transparent models. Motions in a general smooth enough potential near its minimum can be approximately described by a HO. Therefore, if there does exist such effective potential that can equivalently describe the motion of the localized particles relevant to the 2D MITs, the Gaussian form will be probably a most suitable one to be employed. However, considering the success of Mott's original relation and the strong Coulomb interaction present in the samples used for the research of 2D MITs, we are more inclined to the viewpoint that the effective potential felt by the hopping particles are probably relevant to the long-range Coulomb interaction. First, the property of long range ensures that the potential experienced by one particle is the sum of the contributions of many other particles, which is averaged to be smooth and slowly changing. Secondly, the domination of Coulomb interaction over the kinetic energy suppresses the dynamical fluctuations which push the particles away from their most stable position with the lowest potential energy. Therefore it will be a relatively good approximation to represent the effective potential for a hopping particle by an HO potential, with the particle's initial position at the center. For comparison, we give further comments on the traditional form $\exp(-R/\xi)$. This form is specific to a square well potential, which is widely used in models of localization. As we believe, the above form is good for short-range interactions like a hard-core potential, or in a tight bound model where electrons are bound by the ionic potential. However, in case of long-range interactions, the much smoother HO potential is better. This argument may explain why $\nu = 1.5$ is extensively reported in strongly interacting 2D systems, independent of the sample parameters. [16]

Then we discuss what happens when an electrical field E is turned on, and what will be the explicit form of the scaling function with two variables E and T on the insulating side of the MIT. With the assistance of electrical field E, the hopping probability can be written as

$$p(R) \propto \exp(-\alpha\frac{R^2}{\xi^2} - \frac{\Delta}{k_B T} + \frac{ER}{k_B T}). \tag{7}$$

With $z = 1$ in mind, we can define two dimensionless scaling variables:

$$e = E\xi^2, \qquad t = T\xi.$$

Then suppose $R \sim \xi e^y t^x$ (x, y are exponents to be decided), and substitute it into eq(7), we have

$$p(R) \propto \exp[F(e, t)],$$

where $F(e, t)$ is the general scaling function of e and t, and is the sum of the following three parts [17]:

$$\begin{aligned} F_1 &\sim -e^{2y} t^{2x}, \\ F_2 &\sim -e^{-y} t^{-x-1}, \\ F_3 &\sim e^{y+1} t^{x-1}. \end{aligned} \tag{8}$$

We then maximize the eq(7) with respect to R, and get

$$\frac{2\alpha R}{\xi^2} \sim \frac{1}{R^2 T} + \frac{E}{T}.$$

By changing into e and t the above equation can be written as

$$e^y t^x \sim e^{-2y} t^{-2x-1} + e t^{-1}. \tag{9}$$

Therefore the general scaling function is the sum of F_1,F_2 and F_3, with the exponents x and y decided by the constriction equation (9). This suggests that the scaling function will be of different forms for different regions of e and t, and demonstrates complicated behaviors as a result.

We first discuss the simple case of $|e| << 1$. In order for $F(e, t)$ to give a convergent result, we must satisfy:

$$2y \geq 0, \quad -y \geq 0, \quad y + 1 \geq 0.$$

Therefore it is necessary that $y = 0$, which is substituted into eq(9). Then we get $x = -2x - 1$, so $x = -1/3$. We thus arrive at the scaling function for $|e| << 1$,

$$F(e, t) \sim -t^{-2/3} + e t^{-4/3}.$$

For $e = 0$, we obtain

$$\rho(n, T) \propto \exp[-F(0, T)] \propto \exp(\frac{-A\delta n}{T^{2/3}}),$$

which is exactly eq(6) we get earlier.

It is also interesting to understand the behavior of the scaling function for $t << 1$ and $t >> 1$ which will be reported elsewhere because of its complexity.

Before closing, we would like to give more comments on our use of a hopping theory to describe transport properties near MITs. We first point out an important difference of our hopping picture from Mott's original theory. In Mott's formulation, the relevant system is strongly localized, so the localization length ξ is very small and the overlap between adjacent wave-packets is also small. Therefore, in order for a hopping precess to contribute to the conductance, the hopping distance R must be larger than ξ, which puts a constraint on R_{max}. This leads to the conclusion that Mott's hopping is only valid at low enough temperature, or $T < T_0$. However the situation near MITs is quite different, where ξ becomes very large, and the overlap between adjacent wave packets is also large, so a particle hopping between neighboring sites can contribute to the conductance without covering a distance as long as ξ. Therefore the low-temperature-constraint on hopping mechanism is actually absent, and we are justified to apply it inside the critical region $T > T_0$. The relevance of hopping for both $T > T_0$ and $T < T_0$ is supported by inspecting the inset of Fig.2 in ref [3], where no crossover behavior is detected near T_0 for the insulating phase, while the metallic phase shows obvious deviation from the linear behavior for $T < T_0$. In fact, the use of hoppings to study critical phenomena is not new [18], and its relevance to critical exponents (i.e. z) was already established by Efros et al. [14].

In conclusion, we have presented a generalization of Mott's variable range hopping theory to the situation with strong Coulomb interaction in order to understand quantitatively the remarkable discoveries of metal-insulator transitions in two-dimensional systems. In this formulation, the Gaussian form is adopted into the expression of the hopping probability, and the effect of Coulomb gap is also considered. After taking into account the newly proposed scaling consideration, we for the first time give the explicit scaling function of the resistivity on the insulating side, and as a result determine the dynamical and localization length exponents, which are consistent with the experiments. We then clarify the physical meaning contained in our formulation with its relevance to the long-range Coulomb interaction. We also discuss the general scaling function of temperature and electrical field on the insulating side of the transition.

We acknowledge helpful discussions with Z. B. Su and V. Dobrosavljevic. This work is in part supported by National Natural Science Foundation of China.

[1] E. Abrahams, P. W. Anderson, D. C. Licciardello and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).

[2] S.V. Kravchenko *et al.*, Phys. Rev. B**50**, 8039 (1994); D. Simonian, S.V. Kravchenko and M.P Sarachik, preprint cond-mat/9611147v2 (1997); P.T. Coleridge *et al.*, preprint cond-mat/9708118v2 (1997); J. Lam *et al.*, Cond-mat. 9708201 26th August (1997). M. Y. Simmons, A. R. Hamilton, M. Pepper, E. H. Linfield, P. D. Rose and D. A. Rithcie, preprint cond-mat/9709240.

[3] D. Simonian, S. V. Kravchenko and M. P. Sarachik, preprint cond-mat/9611147.

[4] D. Shahar, D. C. Tsui, M. Shayegan, E. Shimshoni and S. L. Sondhi, Science **274**, 589 (1996); D. Shahar, M. Hilke, C. C. Li and D. C. Tsui preprint cond-mat/9706045.

[5] W. J. Zheng, Y. Yu and Z. B. Su, preprint cond-mat/9709225; E. Shimshoni, S. L. Sondhi and D. Shahar preprint cond-mat/9610102.

[6] V. M. Pudalov, preprint cond-mat/9707053, 9707054;

[7] V. Dobrosavljevic, E. Abrahams, E. Miranda and S. Chakravarty, Phys. Rev. Lett. **79**, 455 (1997).

[8] A universal value of $\nu z = 1.6 \mp 0.2$ in Si-based samples has been reported widely, i.e., see: D. Popovic, A. B. Fowler and S. Washburn, Phys. Rev. Lett. **79**, 1543 (1997).

[9] We derive the universality of ν from the general value of νz, because it has been widely accepted that $z \approx 1$ is a universal exponent of 2D strongly interacting systems.

[10] N. F. Mott and E. A. Davis, **Electronic Processes in Non-Crystalline Materials**, Oxford:Clarendon, 2nd edn.(1979).

[11] W. Bayer, PhD thesis, University of Marburg (1974); H. Bottger and V. V. Bryskin, **Hopping Conduction in Solids**, Berlin: Academic Verlag.

[12] A. L. Efros, B. I. Shlovskii, J. Phys. C. **8**, L49 (1975).

[13] T. W. Van Keuls *et al.*, Surface Science **361/362**, 945; W. Mason *et al.*, surface Science **361/362**, 953.

[14] A. L. Efros and M. Pollak, in: Electron Electron Interaction in Disordered Systems, A. L. Efros and M. Pollak eds., North-Holland, Amsterdam, 1985.

[15] For the case of a superconductor-insulator transition, see A. Yazdani and A. Kapitulnik, Phys. Rev. Lett. **74**, 3037 (1995); For the transition between two neighboring quantum Hall plateaus, see. H. P. Wei, L. W. Engel, and D. C. Tsui, Phys. Rev. B **50**, 14 609 (1994).

[16] We presently can not exclude the possibility of a material-dependent ν, since the favorable experiments are all in Si-based material. We also noticed a recent result on GaAs-AlGaAs hole gas by Simmons et al.(cond-mat/9709240), which gave nonuniversal νz. We think this may be due to the presence of strong long-range impurity potential comparable to Coulomb interaction, which invalidates the HO form of the effective potential. Further clarification of this point is highly necessary.

[17] In what follows, we use $\sim$ in equations where constant coefficients are omitted for convenience.

[18] For example, see L. Sheng et al., Phys. Rev. Lett. **79**, 1710 (1997).